AI ACCESSIBILITY HARDWARE

VisionVoiceAssist

Smart glasses that see the world in real time — and describe it out loud, for people who can't see it themselves.

RASPBERRY PI 5 · EDGE AI · REAL-TIME VOICE GUIDANCE

Working hardware prototype · Baku ID 2026

A MOMENT, LIKE ANY OTHER

A blind man stands at a crosswalk.

A car is approaching.

He doesn't know.

“ATTENTION. A car is approaching from your 2 o’clock.”

— VisionVoiceAssist, spoken in real time, before he takes the next step.

This is not a feature. This is the reason VisionVoiceAssist exists.

THE PROBLEM

The world runs on visual information most people never chose to depend on

A white cane or guide dog helps with mobility — but reading a menu, spotting a curb, recognizing a face, or checking whether a crosswalk is clear still requires a sighted person nearby.

2.2B

people live with vision impairment worldwide

43M

are completely blind today — projected to reach 61M by 2050

$411B

lost every year in global productivity due to vision loss

Source: WHO World Report on Vision (2023) · GBD 2020 Global Burden of Disease study

THE GAP

Every existing option solves one piece of the problem — not the whole picture

Cane & Guide Dog

Reliable for mobility and obstacles underfoot.

GAP

No text, faces, or scene detail.

Phone Apps

Seeing AI, Be My Eyes — capable, free, widely used.

GAP

Needs two hands and a raised phone — not hands-free.

Premium AI Glasses

OrCam, Envision — hands-free and accurate.

GAP

$2,000–$4,500, closed hardware, few languages.

VisionVoiceAssist is the first to combine hands-free hardware, real-time scene understanding, offline safety, and an open, affordable build.

OUR SOLUTION

See. Understand. Speak. In under two seconds.

01

SEE

An onboard HD camera continuously captures the world in front of the user.

02

UNDERSTAND

Edge AI and large language models detect obstacles, read text, recognize faces, and interpret the scene.

03

SPEAK

Natural speech describes only what matters — clearly, calmly, and instantly, in the user's own language.

Critical safety alerts run fully offline, on-device. Cloud AI is used only when it adds understanding — never when it costs time.

SYSTEM ARCHITECTURE

The VisionVoiceAssist Brain

1	**Input**	Camera + audio stream from the glasses
↓		
2	**Perception**	Objects, text, faces, and hazards detected
↓		
3	**Memory**	Short-term context — what was just seen
↓		
4	**Reasoning**	LLMs interpret the scene and user intent
↓		
5	**Decision**	Priority & confidence engines choose what matters
↓		
6	**Speech**	Natural, minimal narration generated
↓		
7	**User**	Hears only what's relevant, exactly when it matters

PERCEPTION LAYER — WHAT IT SEES

Four engines turn raw camera frames into structured understanding

1

YOLO

Real-time object detection — obstacles, people, vehicles, doors, signs — at 30+ fps on-device.

2

ByteTrack

Tracks every detected object across frames, so the system knows what's approaching vs. what's still.

3

OCR

Reads menus, labels, signs, and documents the instant they enter the frame.

4

Scene Understanding

A vision-language model turns raw detections into a coherent picture of the environment.

COGNITIVE LAYER — WHAT IT THINKS

Six engines decide what's worth saying — and what isn't

1 Temporal Memory

Remembers what was seen seconds ago, so it won't repeat itself or lose context.

2 Blind Cognitive Layer

Frames every output the way a blind user actually needs it — spatial, relative, actionable.

3 Priority Engine

Ranks hazards above curiosities — a car always outranks a poster.

4 Confidence Engine

Only speaks when detection confidence is high enough to trust.

5 Decision Engine

Chooses whether to alert now, wait, or say nothing at all.

6 Self-Critic

Reviews its own output before speaking, catching contradictions or noise.

SAFETY & PRIVACY LAYER — WHAT IT PROTECTS

Safety never waits on a network connection

1 Offline Safety Core

Hazard, obstacle, and crosswalk detection run entirely on-device — zero dependence on connectivity.

2 Critical Alerts

Danger warnings are never delayed behind a cloud request; they interrupt and speak first.

3 Cloud Fallback

Deeper reasoning (unfamiliar scenes, complex text) can call cloud LLMs only when it's safe to wait.

4 Encryption & Privacy

Visual data is encrypted in transit, processed locally wherever possible, never stored without consent.

USER JOURNEY

A typical day, made a little more independent

Morning

LEAVING HOME

Confirms the door, stairs, and path to the street are clear.

Commute

BUS OR METRO

Announces platform edges, approaching vehicles, and free seats.

Errands

MARKET OR PHARMACY

Reads labels, prices, and dosage instructions aloud, hands-free.

Work / School

SHARED SPACES

Reads documents and whiteboards, recognizes familiar colleagues.

Evening

RETURNING HOME

Guides the way back, flagging anything new or out of place.

FIVE REAL-WORLD SCENARIOS

The same device, five completely different environments

1

Market

Reads product labels and prices, and warns of obstacles in the aisle.

2

Metro

Announces platform edges, approaching trains, and free seats.

3

Street

Detects crosswalks, curbs, and approaching vehicles in real time.

4

Pharmacy

Reads medicine labels and dosage instructions aloud.

5

School

Reads whiteboard text and recognizes classmates, with consent.

LIVE IN ACTION

This isn't a concept. It's a working prototype.

The user picks up a juice carton — VisionVoiceAssist reads the label and describes the object out loud, hands-free, in real time.

1. **Point**
2. **Detect**
3. **Understand**
4. **Speak**

The full loop runs in under 2 seconds — fully on-device, no phone required.

HARDWARE STRATEGY

Why Raspberry Pi 5 — not a custom chip, not a phone

1 Affordable from day one

A ~$80 board keeps the total device under premium competitors' prices by an order of magnitude.

2 Real on-device AI

Enough CPU/GPU headroom to run YOLO, OCR, and tracking locally — no phone, no constant cloud call.

3 Open, repairable, upgradeable

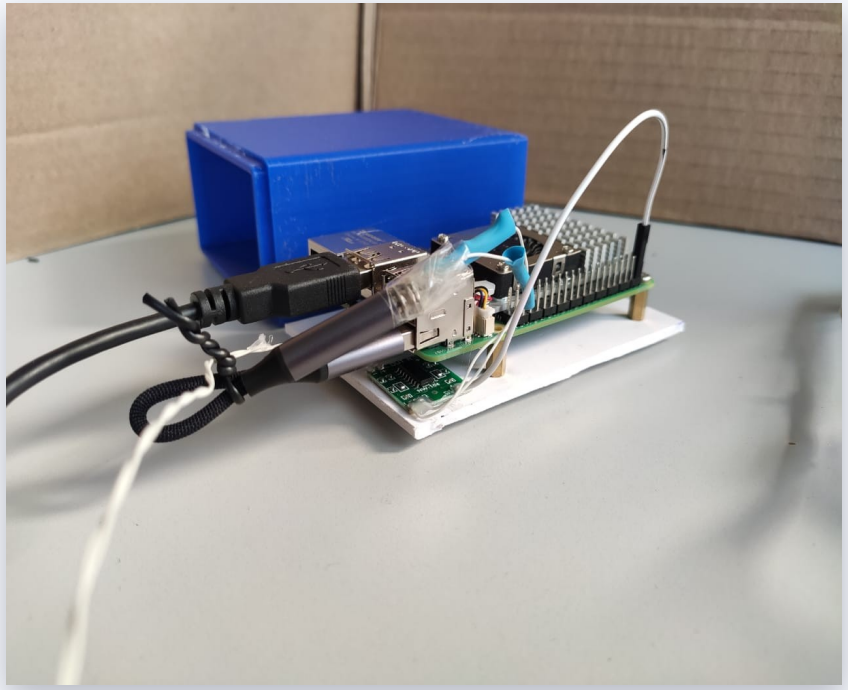
Standard Linux, standard parts. No proprietary lock-in, no single point of failure.

4 A global maker ecosystem

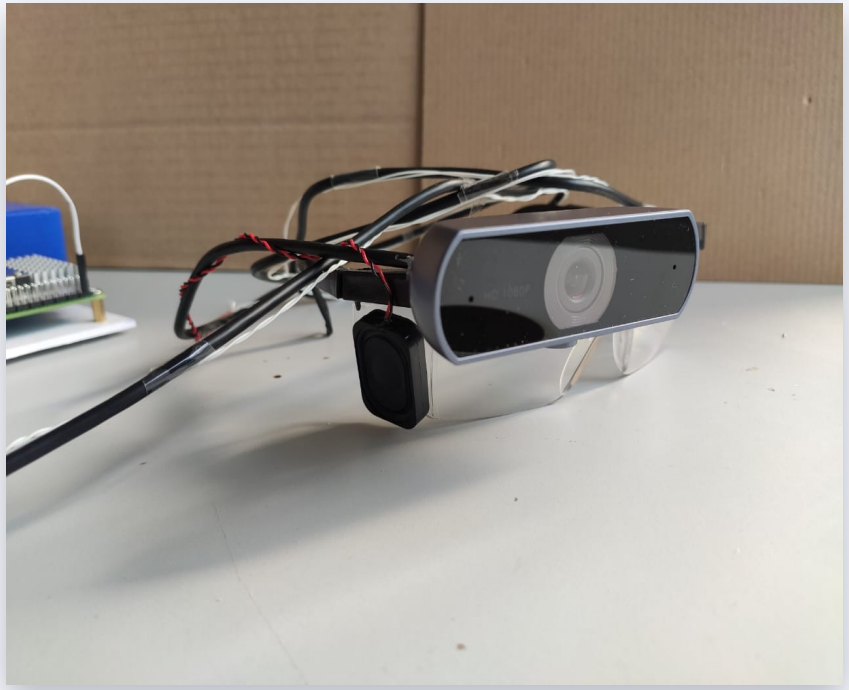
Millions of developers, huge community support, and a clear path to custom future hardware.

HARDWARE

Built on real, accessible components — not vaporware



Raspberry Pi 5 compute unit



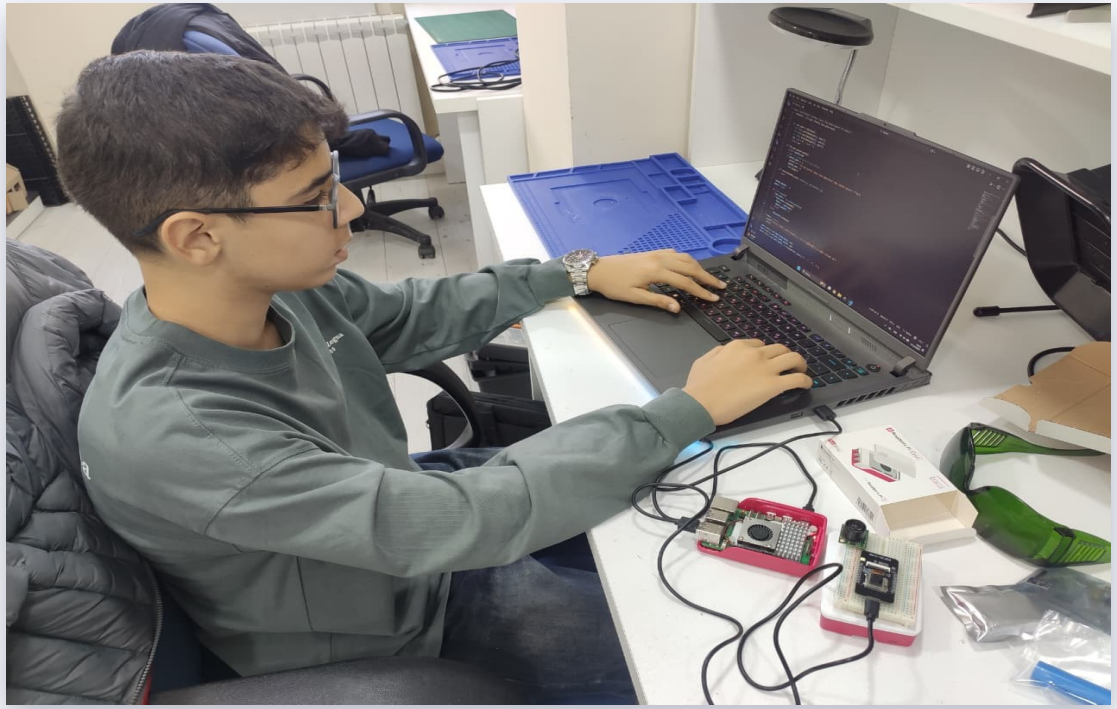
Camera + audio module on the frame

- Raspberry Pi 5 — on-device compute for real-time edge inference
- HD 1080p camera mounted on a lightweight glasses frame
- Compact, discreet speaker for private audio feedback
- Portable battery — several hours of continuous use
- Low-cost, 3D-printable frame — built for local assembly

OUR PROCESS

Every wire and every line of code — designed and tested by our own team

From circuit simulation to Python vision pipelines, VisionVoiceAssist was built end-to-end in-house — hardware assembly, embedded software, and AI integration, iterated by hand through dozens of prototype cycles.



Coding the detection pipeline on Raspberry Pi 5

Prototyping the power circuit before the physical build

FOUNDER STORY

“Why can’t AI become the eyes of someone who cannot see?”

VisionVoiceAssist started from that simple question. It led to hundreds of hours of engineering, hardware prototyping, and AI development — turning a question into a working system.

Every prototype in this deck — the wiring, the code, the trained pipelines — was built by hand, one iteration at a time.

TEAM

Two co-founders, one shared mission

Aliasgar Fatullayev

Founder & Lead AI Engineer

Responsible for the end-to-end design and development of VisionVoiceAssist, including AI architecture, computer vision, embedded systems, edge computing, and product engineering.

CORE EXPERTISE

- AI & Computer Vision
- Embedded Systems (Raspberry Pi)
- Edge AI & Real-Time Inference
- System Architecture
- Cybersecurity & Privacy

Aliakber Fatullayev

Co-Founder & Marketing

Leads outreach, storytelling, and early community engagement — building relationships with schools, families, and accessibility organizations, and shaping how VisionVoiceAssist is presented to the world.

CORE EXPERTISE

- Brand & Storytelling
- Community Outreach
- Pilot & Partner Coordination
- Content & Social Media

Both founders currently dedicate 5–6 hours daily to VisionVoiceAssist alongside their studies.

VALIDATION

Where the project actually stands today

Current Validation Status

✅ Functional hardware prototype completed

✅ Core AI pipeline operating in real time

✅ End-to-end AI workflow successfully demonstrated

✅ 14 informal testers provided hands-on feedback

🔲 Structured pilot with visually impaired users planned

Early testers flagged unclear audio output — already identified and resolved before this deck was finalized.

Performance Goals

Rather than claim audited metrics this early, we're optimizing toward:

- Sub-2 second response time
- Real-time scene understanding
- Low-latency voice feedback
- High object detection reliability
- Offline-first safety processing
- Target accuracy above 95% (to be validated in pilot)

KEY FEATURES

Six capabilities, one continuous experience

01 Obstacle & Hazard Detection

Stairs, curbs, holes, crosswalks, and approaching vehicles — flagged before they become a risk.

02 Text & Document Reading

OCR reads menus, labels, signs, and documents aloud on demand.

03 Familiar Face Recognition

Identifies people the user has introduced to the system, with consent.

04 Natural Scene Narration

Describes surroundings the way a friend would — only what's relevant.

05 Offline Safety Mode

Critical hazard detection runs fully on-device, no connectivity required.

06 Privacy By Design

Visual data is encrypted and processed locally wherever possible.

PROTOTYPE METRICS

What we're building toward, tracked from day one

Response Time

Sub-2 second, point-to-speech loop

Frame Rate

30+ fps object detection on-device

Battery

Multi-hour continuous use, current pack

Accuracy

Target 95%+, validated during pilot

Offline Detection

100% of safety alerts run without network

Latency Source

Edge-first architecture, cloud only when safe to wait

Targets, not audited results — VisionVoiceAssist is a hardware prototype tracking toward these goals through pilot testing.

COMPETITIVE LANDSCAPE

Fourteen capabilities that define this category — head to head

Capability	OrCam MyEye	Envision Glasses	Google Lookout	Seeing AI	Ray-Ban Meta	VisionVoice Assist
Hands-free glasses form factor	✓	✓	–	–	✓	✓
Continuous real-time narration	–	–	–	–	–	✓
Obstacle & hazard detection	–	–	–	–	–	✓
Crosswalk & traffic detection	–	–	–	–	–	✓
OCR / text reading	✓	✓	✓	✓	✓	✓
Face recognition (consented)	✓	✓	–	✓	–	✓
Fully offline safety alerts	✓	–	✓	–	–	✓
Azerbaijani language support	–	–	–	–	–	✓
Open / upgradeable hardware	–	–	–	–	–	✓
On-device AI, no phone needed	✓	–	–	–	–	✓
LLM-based scene reasoning	✓	✓	✓	✓	✓	✓
Privacy-first local processing	✓	–	✓	–	–	✓
Price under $500	–	–	✓	✓	✓	✓
Purpose-built for blind users	✓	✓	✓	✓	–	✓

Based on publicly available product information as of 2026. Features vary by hardware version, region, and subscription tier.

MARKET OPPORTUNITY

A large, underserved, and fast-growing market

$6.3B → $11.2B

Global assistive-technology market for the visually impaired, 2025 → 2030 (est., ~12% CAGR)

2.2B

potential users worldwide living with vision impairment

Why now

- Edge AI hardware (like Raspberry Pi 5) finally makes real-time on-device inference affordable.
- Aging populations are driving demand for assistive technology worldwide.
- Accessibility regulation is expanding across public and private sectors.

Estimates based on industry market research (Mordor Intelligence, 2025) and WHO population data. Directional planning assumptions, not guaranteed outcomes.

TIMING

Why Now?

1 Edge AI has arrived

Hardware is finally powerful enough for real-time wearable intelligence.

2 LLMs became practical

Large vision models have reached real-world usability, not just lab demos.

3 Demand keeps growing

Global demand for accessible technology continues to rise every year.

4 Hardware got affordable

Capabilities once locked behind expensive proprietary devices are now reachable.

VisionVoiceAssist combines these shifts into one practical product for independent mobility.

MARKET DEMAND

Who this is for — and who we're building it with

- **Blind Organizations**
- **Schools**
- **NGOs**
- **Rehabilitation Centers**
- **Hospitals**

Early, informal conversations with a teacher and a parent of prospective users confirmed genuine interest in a pilot. We are now preparing structured pilot programs with accessibility organizations, schools, and rehabilitation centers to build on this initial validation.

BUSINESS MODEL

Hardware to start. Recurring revenue as we scale.

1

Device Sales

One-time hardware purchase, priced well below premium competitors.

2

Government & NGO Programs

Bulk procurement for social services, blind schools, and rehab centers.

3

Cloud AI Subscription

Optional upgrade for advanced multilingual reasoning and cloud features.

4

Enterprise & Accessibility Partners

Retail, transit, and public-space integrations for accessibility compliance.

FUNDING STRATEGY

How we will use $100,000

Investment Area	Allocation
Hardware Engineering & Prototype Iteration	30%
AI & Edge Computing Development	30%
Pilot Programs & User Validation	20%
Manufacturing Preparation	10%
Certification, Legal & Operations	10%

This investment will enable VisionVoiceAssist to transition from a validated prototype into a pilot-ready assistive technology platform, preparing the product for manufacturing and broader market adoption.

RISK ANALYSIS

What could go wrong — and how we're already addressing it

RISK

Battery life under real use

↓

MITIGATION

Ongoing firmware & power-management optimization.

RISK

Latency in complex scenes

↓

MITIGATION

Edge-first AI architecture keeps critical inference on-device.

RISK

Recruiting pilot participants

↓

MITIGATION

Structured outreach via blind-community organizations, incl. Blind Union.

RISK

Manufacturing cost at scale

↓

MITIGATION

Phased manufacturing partnerships once pilot demand is validated.

ROADMAP

From working prototype to global reach

1

2026

WORKING PROTOTYPE

→

2

Pilot

NGOS & SCHOOLS

→

3

100

USERS ONBOARDED

→

4

1,000

USERS ONBOARDED

5

Azerbaijan

NATIONAL COVERAGE

→

6

Caucasus

REGIONAL EXPANSION

→

7

Europe

MARKET ENTRY

→

8

World

GLOBAL PLATFORM

Milestones represent our planning targets, not guaranteed timelines or committed dates.

FUTURE VISION

What VisionVoiceAssist becomes after the pilot

1 World Model

A persistent spatial memory of places the user visits often — home, school, work — for faster, calmer guidance.

2 Developer SDK

An open API so hospitals, schools, and NGOs can build custom accessibility tools on top of the platform.

3 Expanded Languages

Beyond Azerbaijani and English — regional languages across the Caucasus and Middle East.

4 Next-Gen Hardware

A lighter, custom PCB successor to the Raspberry Pi prototype, purpose-built for daily wear.

OUR VISION

Not just smart glasses.
A full AI decision system built to help people understand the world around them.

We're not building a gadget. We're building independence — one glance, one sentence, one step at a time.

Thank you.